

Wenn "mehr" nicht gleichbedeutend mit "besser" ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen

Koch, Achim

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Koch, A. (1998). Wenn "mehr" nicht gleichbedeutend mit "besser" ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. *ZUMA Nachrichten*, 22(42), 66-90. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-208299>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

WENN „MEHR“ NICHT GLEICHBEDEUTEND MIT „BESSER“ IST: AUSSCHÖPFUNGSQUOTEN UND STICHPROBENVERZERRUNGEN IN ALLGEMEINEN BEVÖLKERUNGSUMFRAGEN

ACHIM KOCH

Wenn Ausfälle bei der Realisierung von Zufallsstichproben nicht zufällig erfolgen, gefährden sie die Generalisierbarkeit von Umfrageergebnissen. In der Praxis wird häufig angenommen, daß eine Stichprobe um so weniger verzerrt ist, je höher die Ausschöpfung der Umfrage ausfällt. In dem Aufsatz wird auf Basis von sechs allgemeinen Bevölkerungsumfragen, deren Ausschöpfung zwischen 50 und 80 Prozent variiert, analysiert, inwiefern diese Annahme gerechtfertigt ist. Ein Vergleich der Randverteilungen ausgewählter soziodemographischer Merkmale mit Daten des Mikrozensus ergibt keine empirischen Belege dafür, daß die Stichprobenverzerrungen der Umfragen mit höherer Ausschöpfung geringer sind. Als eine mögliche Erklärung für diesen Befund wird diskutiert, inwiefern die von Umfrageinstituten berichteten Ausschöpfungsquoten zuverlässig sind.

Nonresponses in sample surveys jeopardize the generalizability of survey results, if they do not occur completely randomly. It is often assumed that the extent of nonresponse bias increases in line with nonresponse rates. The article discusses data from six general population surveys with response rates varying between 50 and 80 percent to see whether this, in fact, is the case. The marginals of several sociodemographic variables were compared with official microcensus data. No empirical evidence was found of a straightforward relationship between response rates and the extent of nonresponse bias. The reliability of survey organization reports of response rates is discussed as a possible explanation.

1. Einleitung

Bei Umfragen, die auf Zufallsstichproben basieren, gelingt es so gut wie nie, mit allen ausgewählten Zielpersonen ein Interview durchzuführen. Ein Teil der Personen wird während der Feldzeit nicht erreicht, ein Teil wird zwar kontaktiert, ist aber nicht fähig oder nicht bereit, an der Umfrage teilzunehmen. Wenn sich diese Ausfälle nicht zufällig über die Zielpopulation verteilen, liegt eine Stichprobenverzerrung bzw. ein Non-response-Bias vor, durch den die Generalisierbarkeit der Umfrageergebnisse beeinträchtigt wird.

In der Praxis wird häufig die Ausschöpfungsquote als Proxy für eine Messung des Non-response-Bias herangezogen. Man nimmt dabei an, daß die Verzerrung um so geringer ist, je höher die Teilnahmequote ausfällt.¹⁾ Dies ist aber nur begrenzt richtig, denn die Ausschöpfungsquote stellt lediglich eine Komponente für die Stärke des Nonresponse-Bias dar.

Bei linearen Statistiken wie Mittelwerten oder Anteilswerten ist der Nonresponse-Bias eine multiplikative Funktion der Höhe der Ausfallquote einerseits und der Differenz zwischen Teilnehmern und Nichtteilnehmern der Umfrage bei der betrachteten Variablen andererseits (vgl. Groves/Couper 1992: 103).

Der Nonresponse-Bias eines Anteilswertes z.B. berechnet sich nach

$$\text{Ausfallquote} * (p_T - p_{NT}),$$

wobei p_T dem Anteilswert der Teilnehmer und p_{NT} dem Anteilswert der Nichtteilnehmer bei der betreffenden Variablen entspricht. Die Formel verdeutlicht zweierlei:

1. Aus der Höhe der Ausfallquote kann nur auf das mögliche *Potential* eines Non-response-Bias geschlossen werden: je geringer die Ausschöpfung, desto größer kann die Verzerrung ausfallen. Prinzipiell ist es jedoch auch bei einer nur geringen Ausschöpfung möglich, daß der Bias klein ist, nämlich dann, wenn die Ausfälle weitgehend zufällig erfolgen, dh. wenn die Unterschiede zwischen Teilnehmern und Nichtteilnehmern bei einer Variablen gering sind.

2. Der Nonresponse-Bias ist eine *merkmalspezifische* Größe. Eine schlecht ausgeschöpfte Umfrage kann bei einem Merkmal eine deutliche Verzerrung aufweisen (weil sich Teilnehmer und Nichtteilnehmer hinsichtlich der betreffenden Variablen unterscheiden), bei einem anderen Merkmal jedoch weitgehend unverzerrt sein (weil Teilnehmer und Nichtteilnehmer bezüglich dieser Variablen nicht oder nur wenig differieren).

Will man der Frage des Nonresponse-Bias empirisch nachgehen, steht man vor dem Problem, daß die beiden Bestandteile dieser Verknüpfung in unterschiedlichem Maße einer Messung zugänglich sind: während die Ausfallquote - zumindest theoretisch - gut und eindeutig ermittelbar ist, liegen Informationen über die zweite Komponente - den Unterschied zwischen Teilnehmern und Nichtteilnehmern - normalerweise nicht vor. Die Abwesenheit von Informationen über die Nichtteilnehmer ist die Ursache dafür, daß eine routinemäßige Messung des Nonresponse-Bias kaum möglich ist.

Es gibt verschiedene Möglichkeiten, sich dem Problem der Ermittlung des Nonresponse-Bias zu nähern.²⁾ Eine häufig verfolgte Methode ist der Vergleich von Umfrageergebnissen mit Informationen der amtlichen Statistik. Bei Erhebungen wie der Volkszählung oder dem Mikrozensus spielt das Problem des Nonresponse aufgrund der Teilnahme-pflicht eine geringe Rolle. Im Mikrozensus etwa liegt der Unit-Nonresponse bei ca. drei Prozent (vgl. Lüttinger/Riede 1997: 27). Man kann deshalb annehmen, daß die Ergebnisse des Mikrozensus ein einigermaßen valides Außenkriterium zur Beurteilung möglicher Stichprobenverzerrungen von Umfragedaten darstellen.³⁾

Ich werde im folgenden die Randverteilungen ausgewählter Merkmale des Mikrozensus mit Daten aus sechs allgemeinen Bevölkerungsumfragen vergleichen, deren berichtete Ausschöpfung zwischen 50 und 80 Prozent beträgt. Aufgrund der begrenzten Zahl der im Mikrozensus verfügbaren Merkmale muß sich der Vergleich auf einige soziodemographische Variablen beschränken. Leitfragen der Analyse sind:

1. Welche Personengruppen sind in allgemeinen Bevölkerungsumfragen unter- bzw. überrepräsentiert?
2. Ist die Richtung der Verteilungsabweichungen in den betrachteten Umfragen ähnlich?
3. Gibt es empirische Anhaltspunkte dafür, daß der Nonresponse-Bias in Umfragen mit höherer Ausschöpfung niedriger ist als in solchen mit schlechterer Ausschöpfung?

2. Die analysierten Umfragen

Die Datenbasis bilden sechs allgemeine Bevölkerungsumfragen, die alle für sich in Anspruch nehmen, hohen Qualitätsstandards zu genügen. Fünf der betrachteten Umfragen sind sozialwissenschaftliche Surveys, die von ZUMA betreut wurden. Es sind dies der ALLBUS 1992 und 1994, der Wohlfahrtssurvey 1993 sowie zwei Umfragen des Sozialwissenschaften-Bus aus dem Jahre 1993. Berücksichtigt wird außerdem die Presse-medientranche der Media-Analyse 1994.⁴⁾ Alle sechs Surveys wurden in Form persön-

lich-mündlicher Interviews durchgeführt, die durchschnittliche Interviewdauer lag zwischen 50 (Media-Analyse) und 80 Minuten (ALLBUS 1992). In Übersicht 1 sind die wichtigsten Informationen zu Grundgesamtheit, Stichprobe und Ausschöpfung der Umfragen zusammengestellt.

Die Definition der Grundgesamtheit variiert etwas zwischen den einzelnen Umfragen. Alle sechs Surveys beziehen sich auf die Bevölkerung in Privathaushalten; Personen, die in Anstalten leben, wurden nicht befragt. Der Wohlfahrtssurvey und die beiden Umfragen des Sozialwissenschaften-Bus beschränken sich - wie die meisten sozialwissenschaftlichen Umfragen - auf die erwachsene deutsche Bevölkerung in Privathaushalten. In den beiden ALLBUS-Erhebungen dagegen wurden zusätzlich zur deutschen Bevölkerung auch die in Deutschland lebenden Ausländer befragt - dies allerdings nur dann, wenn das Interview in deutscher Sprache durchgeführt werden konnte. In der Media-Analyse wurden nicht nur erwachsene Deutsche, sondern auch Jugendliche zwischen 14 und 17 Jahren interviewt.

In fünf der sechs betrachteten Umfragen wurde das dreistufige Stichprobendesign des Arbeitskreises Deutscher Marktforschungsinstitute („ADM-Stichprobensystem“) eingesetzt, das in der Bundesrepublik bei der Ziehung von Zufallsstichproben für persönlich-mündliche Bevölkerungsumfragen eine dominierende Stellung einnimmt.⁵⁾ Bei diesem Verfahren wird in der ersten Stufe eine Stichprobe von Wahlbezirken gezogen. In der zweiten Auswahlstufe ermitteln die Interviewer in den ausgewählten Wahlbezirken anhand eines zufällig ausgewählten Startpunktes und vorgegebener Begehungsregeln eine festgelegte Zahl von Haushalten.⁶⁾ In der dritten Ziehungsstufe wählen die Interviewer mit Hilfe eines vorbereiteten Zufallsschemas („Schwedenschlüssel“) die zu befragende Person aus den zur Grundgesamtheit der Umfrage zählenden Haushaltsmitgliedern aus.⁷⁾

Eine Alternative zum ADM-Design, die in der Praxis allerdings nur selten gewählt wird, stellt die im ALLBUS 1994 erfolgte Stichprobenziehung aus Einwohnermelderegistern dar. Bei diesem zweistufigen Verfahren wird zunächst eine Stichprobe von Gemeinden gebildet; im ALLBUS 1994 waren es 151 Gemeinden. In der zweiten Auswahlstufe wurden anschließend aus den Einwohnermelderegistern dieser Gemeinden jeweils 40 Personen für die Befragung zufällig ausgewählt. Die Interviewer erhielten die zu befragenden Personen mit Namen und Adresse vorgegeben. Die Phase der Stichprobenbildung ist bei diesem Verfahren damit vollständig von der Feldarbeit getrennt.

Übersicht 1: Die analysierten Umfragen

	ALLBUS 1992	ALLBUS 1994	Wohlfahrts- survey 1993	Sozialwissen- schaften-Bus III 1993*	Sozialwissen- schaften-Bus II 1993*	Media-Analyse (Pressemedientranche) 1994
Umfrageinstitut	Infratest			GFM-GETAS		Basis Research, GFM-GETAS, Ifak, Marplan, Media Markt Analysen, Sample
Grundgesamtheit	Erwachsene Bevölkerung in Privathaushalten		Erwachsene deutsche Bevölkerung in Privathaushalten			Deutsche Bevölkerung in Privathaushalten im Alter von 14 und mehr Jahren
Stichprobenverfahren	ADM-Stichprobe „Standard-Random“	Einwohnermelde- register- Stichprobe	ADM-Stichprobe „Random-Route“			ADM-Stichprobe „Adress-Random“
Netto-Fallzahl						
West	2400	2342	2046	2024	2143	19247
Ost	1148	1108	1016	2142	1092	6374
Ausschöpfung (in %)						
West	51,9	53,2	62,9	68,0	73,0	80,2
Ost	54,7	55,2	64,4	71,1	72,9	80,2

* Der Sozialwissenschaften-Bus I 1993 wurde nicht berücksichtigt, weil er lediglich in Ostdeutschland durchgeführt wurde.

Sowohl das ADM-Design als auch der Zugang über die Einwohnermelderegister führen zu Zufallsstichproben der Bevölkerung: der Unterschied besteht allerdings darin, daß bei der Einwohnermelderegister-Stichprobe die Befragungspersonen eine gleiche Auswahlchance haben, während dies beim ADM-Design nur für die Befragungshaushalte gilt. Die Auswahlchance der einzelnen Haushaltsmitglieder ist hier um so geringer, je größer die Zahl der Haushaltspersonen ist, die zur Grundgesamtheit der betreffenden Umfrage zählen. Dies hat zur Folge, daß bei Auswertungen auf Personenebene vom Design her eine Gewichtung mit der - um die nicht zur Grundgesamtheit zählenden Haushaltsmitglieder reduzierten - Haushaltsgröße vorgesehen ist (exakter: mit einem Faktor, der dem reziproken Wert der Auswahlchance entspricht).

In den fünf sozialwissenschaftlichen Umfragen wurden jeweils ungefähr 2.000 bis 2.400 Personen in Westdeutschland und 1.000 Personen in Ostdeutschland befragt. Lediglich im Sozialwissenschaften-Bus III 1993 wurden auch im Osten ca. 2.000 Fälle realisiert. Die Media-Analyse dagegen weist deutlich höhere Fallzahlen auf: in dieser Gemeinschaftsuntersuchung, an deren Durchführung insgesamt sechs Umfrageinstitute beteiligt waren, wurden rund 19.000 Fälle im Westen und 6.000 im Osten realisiert. Jedes beteiligte Umfrageinstitut hat ca. 3.000 Interviews in Westdeutschland und 1.000 in Ostdeutschland durchgeführt.

In den Ausschöpfungsquoten der Studien, wie sie von den Umfrageinstituten berichtet werden, bestehen große Unterschiede.⁸⁾ Die beiden ALLBUS-Umfragen weisen mit etwas über 50 Prozent die niedrigste Ausschöpfung auf, es folgt der Wohlfahrtssurvey mit etwas über 60 Prozent, dann die beiden Umfragen des Sozialwissenschaften-Bus mit ungefähr 70 Prozent. Spitzenreiter ist die Media-Analyse mit 80 Prozent. West- und Ostdeutschland unterscheiden sich bei den einzelnen Umfragen kaum in der Ausschöpfung.⁹⁾

3. Vergleich von Umfrageergebnissen mit Mikrozensusdaten

Um zu untersuchen, ob diese unterschiedlichen Ausschöpfungsquoten einen Einfluß auf den Nonresponse-Bias der Umfragen haben, werden im folgenden die Randverteilungen ausgewählter demographischer Merkmale mit Ergebnissen des Mikrozensus verglichen. Der Mikrozensus ist eine seit 1957 (nahezu) jedes Jahr stattfindende amtliche Erhebung, bei der die Erfassung erwerbsstatistischer und soziodemographischer Merkmale im Mittelpunkt steht. Der Erhebung liegt eine Ein-Prozent-Haushaltsstichprobe der wohnberechtigten Bevölkerung zugrunde. Sie umfaßt seit 1991 etwa 730.000 befragte Personen in 330.000 Haushalten im wiedervereinigten Deutschland. Für die ausgewählten Personen und Haushalte besteht Teilnahmepflicht (vgl. Lüttinger/Riede 1997).

Die vorliegende Analyse basiert auf der ZUMA zur Verfügung stehenden 70-Prozent-Unterstichprobe des Mikrozensus 1993. Mit der Wahl dieses Mikrozensus nehme ich in Kauf, daß die Erhebungszeitpunkte der beiden ALLBUS-Umfragen jeweils um ein Jahr abweichen.¹⁰⁾ Da sich die Verteilungen sozialstruktureller Merkmale normalerweise nur langsam ändern, erscheint dies vertretbar.

In den Vergleich einbezogen wurden soziodemographische Merkmale, für die bereits in früheren Untersuchungen Verteilungsabweichungen zwischen Mikrozensus und Umfragedaten festgestellt wurden. Hartmann/Schimpl-Neimanns (1992) fanden bei einem Vergleich der ALLBUS-Erhebungen 1986, 1988 und 1990 mit Daten des Mikrozensus, daß in den Umfragedaten Frauen, über 70jährige, Personen mit Hauptschulabschluß, Erwerbstätige, Arbeiter, Verwitwete, Geschiedene und Ledige sowie Alleinlebende unterrepräsentiert sind (vgl. Übersicht 2).¹¹⁾ In der Übersicht kursiv dargestellt sind dabei Ergebnisse, die als relativ gut belegt gelten können, weil sie auch in anderen Untersuchungen zur „Demographie“ von Ausfällen wiederholt beobachtet wurden (vgl. z.B. die Zusammenfassung empirischer Befunde bei Burton 1996: 13ff., Esser u.a. 1989: 111ff., Goyder 1987: 80ff., Groves 1989: 201ff.). Dies betrifft die Untererfassung von Älteren, Personen mit niedrigem Bildungsabschluß sowie Arbeitern (These des Mittelschichtbias) und alleinlebenden Personen.

Übersicht 2: Untersuchte Merkmale

Merkmal	Unterrepräsentiert sind
Geschlecht	Frauen
Alter	<i>über 70jährige</i>
Schulabschluß	<i>Personen mit Hauptschulabschluß</i>
Stellung im Erwerbsleben	Erwerbstätige
Berufliche Stellung	<i>Arbeiter</i>
Familienstand	Verwitwete, Geschiedene, Ledige
Haushaltsgröße	<i>Alleinlebende</i>

Damit Verteilungsabweichungen zwischen amtlicher Statistik und Umfragedaten als Ausdruck eines Nonresponse-Bias interpretiert werden dürfen, müssen die zugrundeliegenden Zielpopulationen und Erhebungskonzepte vergleichbar sein. Um die Bevölkerungskonzepte zu vereinheitlichen, basieren die folgenden Auswertungen auf der erwachsenen deutschen Bevölkerung in Privathaushalten. Zur Erzielung einer weitgehend äquivalenten Operationalisierung der Merkmale wurden die Variablen jeweils auf einen gemeinsamen Standard rekodiert. Die vorgenommenen Rekodierungen für die Merkmale Schulabschluß, Stellung im Erwerbsleben und berufliche Stellung, die

etwas komplexeren Charakter haben, sind im Anhang dokumentiert. Um zu vermeiden, daß zufällige Differenzen aufgrund des Stichprobenfehlers fälschlicherweise als Folge eines Nonresponse-Fehlers interpretiert werden, werde ich im folgenden nur Abweichungen, die bei allen (oder nahezu allen) Umfragen *in gleicher Richtung* auftreten, als Beleg für eine Stichprobenverzerrung werten.¹²⁾ Eine zusätzliche Absicherung erfährt diese Interpretation immer dann, wenn die Abweichungen mit dem gegebenen Vorwissen über die Systematik von Ausfällen übereinstimmen.

4. Ausschöpfungsquoten und Stichprobenverzerrungen

Die folgenden Tabellen enthalten jeweils die Merkmalsverteilung für den Mikrozensus. Für die einzelnen Surveys werden die Prozentsatzdifferenzen zum Mikrozensus wiedergegeben. Negative Abweichungen sind zur besseren Orientierung kursiv dargestellt. Die Umfragen sind nach der Ausschöpfung in aufsteigender Reihenfolge angeordnet.

Sämtliche Tabellen basieren auf der erwachsenen deutschen Bevölkerung in Privathaushalten. Für die Studien nach dem ADM-Verfahren wurden die designgewichteten Daten verwendet.¹³⁾

Geschlecht

In Westdeutschland stimmt die Verteilung nach dem Merkmal Geschlecht in fast allen Umfragen mit den Daten des Mikrozensus recht gut überein. Eine Ausnahme bildet lediglich der ALLBUS 1994, bei dem der Anteil der Frauen um 3,8 Prozentpunkte unter dem Wert des Mikrozensus liegt.¹⁴⁾

Tabelle 1: Differenz zum Mikrozensus beim Merkmal Geschlecht (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Männlich	47,2	1,8	3,8	0,2	-1,0	1,3	-0,3
Weiblich	52,8	-1,8	-3,8	-0,2	1,0	-1,3	0,3
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Männlich	46,7	2,2	1,4	4,7	2,4	3,9	1,5
Weiblich	53,3	-2,2	-1,4	-4,7	-2,4	-3,9	-1,5

In Ostdeutschland sind beim Geschlecht die Differenzen zum Mikrozensus etwas größer als in Westdeutschland. In allen sechs Umfragen liegt der Anteil der Frauen unter dem Ergebnis des Mikrozensus, am stärksten ist die Abweichung im Wohlfahrtssurvey mit - 4,7 Prozentpunkten, am geringsten beim ALLBUS 1994 mit -1,4 Prozentpunkten. Die Gleichförmigkeit der Abweichung über alle sechs Umfragen stellt ein Indiz für einen - schwach ausgeprägten - Nonresponse-Bias dar.¹⁵⁾ Hinweise darauf, daß die Verzerrung mit steigender Ausschöpfung geringer wird, finden sich nicht.

Alter

In Westdeutschland ist die Unterrepräsentierung der über 70jährigen das hervorstechende Ergebnis bei der Altersverteilung. In allen sechs Umfragen ist der Anteil dieser Altersgruppe um 3,2 bis 4,5 Prozentpunkte niedriger als im Mikrozensus. Im Gegenzug sind in allen Umfragen die 30- bis 49jährigen stärker als im Mikrozensus vertreten. Der überproportionale Ausfall älterer Personen steht in Einklang mit den bekannten Regelmäßigkeiten des Teilnahmeverhaltens, man kann hier also mit großer Wahrscheinlichkeit annehmen, daß ein Nonresponse-Effekt vorliegt. Eine bessere Anpassung an den Mikrozensus mit zunehmender Ausschöpfung der Umfragen ist nicht festzustellen.

Tabelle 2: Differenz zum Mikrozensus beim Merkmal Alter (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
18 bis 29 Jahre	20,8	0,4	0,2	-2,6	-0,5	1,7	-1,6
30 bis 39 Jahre	18,1	3,0	2,7	2,0	2,9	1,1	1,6
40 bis 49 Jahre	15,3	1,9	0,4	1,1	1,3	2,4	2,7
50 bis 59 Jahre	18,3	-0,2	0,5	1,8	-1,0	-1,0	0,3
60 bis 69 Jahre	13,8	-0,5	-0,3	1,4	0,8	0,4	0,2
70 Jahre und älter	13,7	-4,5	-3,6	-3,7	-3,5	-4,5	-3,2
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
18 bis 29 Jahre	19,7	-1,7	-1,7	-7,2	-5,2	0,1	-4,0
30 bis 39 Jahre	19,9	2,8	-1,5	0,8	1,0	-1,1	1,1
40 bis 49 Jahre	15,9	2,2	2,2	2,3	2,2	3,2	2,7

50 bis 59 Jahre	19,4	3,5	1,6	3,8	2,3	0,8	1,8
60 bis 69 Jahre	13,3	-1,4	0,7	4,1	3,4	1,8	3,1
70 Jahre und älter	11,8	-5,4	-1,3	-3,8	-3,8	-4,9	-4,7

Auch in Ostdeutschland ist in allen Umfragen eine Unterrepräsentierung der über 70jährigen zu beobachten. Diese fällt zumeist in ähnlicher Größenordnung aus wie in Westdeutschland.¹⁶⁾ In drei Umfragen ist darüber hinaus die jüngste Altersgruppe der 18- bis 29jährigen deutlich zu gering vertreten: dies betrifft die Media-Analyse (-4,0 Prozentpunkte), den Sozialwissenschaften-Bus III 1993 (-5,2 Prozentpunkte) und den Wohlfahrtssurvey 1993 (-7,2 Prozentpunkte). Die Stärke der Abweichung und die z.T. auch in anderen Studien berichtete höhere Ausfallquote jüngerer Altersgruppen (vgl. z.B. Burton 1996: 15) legen es nahe, auch diese Abweichung als systematische Stichprobenverzerrung zu werten. Im Vergleich zum Mikrozensus überrepräsentiert sind im Osten in (nahezu) allen Umfragen die 40- bis 69jährigen. Insgesamt betrachtet weist der ALLBUS 1994 mit lediglich 55 Prozent Ausschöpfung die beste Übereinstimmung mit dem Mikrozensus auf, bei allen anderen Umfragen sind die Abweichungen deutlich größer.¹⁷⁾

Schulabschluß

Beim Merkmal Schulabschluß ist nach der Literatur eine Unterrepräsentierung von Personen mit niedrigem Bildungsabschluß zu erwarten. In Westdeutschland trifft dies für alle sechs Umfragen zu: der Anteil der Personen mit Volks-/Hauptschulabschluß liegt konsistent unter dem Wert des Mikrozensus.¹⁸⁾ Auf der anderen Seite sind Personen mit mittleren Bildungsabschlüssen und - in etwas schwächerem Maße - auch Personen mit Abitur oder Hochschulbildung (Ausnahme: Media-Analyse) in den Umfragen überrepräsentiert. Den größten Unterschied zum Mikrozensus weist die Umfrage mit der zweithöchsten Ausschöpfung auf: im Sozialwissenschaften-Bus II 1993 liegt der Anteil der Volks- und Hauptschulabsolventen um 10,3 Prozentpunkte unter dem Wert des Mikrozensus. Am besten stimmt die Media-Analyse mit den Daten der amtlichen Statistik überein.

In Ostdeutschland kann aufgrund der Unterschiede in den Abfragetechniken lediglich zwischen Abitur/Fachhochschule/Hochschule auf der einen Seite und allen niedrigeren Abschlüssen auf der anderen Seite differenziert werden (vgl. Anhang).¹⁹⁾ Bei dieser Differenzierung zeigt sich eine systematische Überrepräsentierung von Personen mit Abitur bzw. Fachhoch-/Hochschulabschluß, während Personen mit niedrigeren Abschlüssen in allen Umfragen zu gering vertreten sind. Die Differenzen zum Mikrozensus sind mit 4,3 bis 6,3 Prozentpunkten annähernd gleich groß, eine Abnahme der Verzerrung mit zunehmender Ausschöpfung der Umfragen ist nicht festzustellen.

Tabelle 3: Differenz zum Mikrozensus beim Merkmal Schulabschluß (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Bis Volks-/Hauptschul- abschluß	59,6	-6,2	-5,7	-7,6	-6,0	-10,3	-3,5
Mittlere Reife, Fachhochschulreife	23,5	3,7	4,2	5,1	5,1	7,3	3,8
Abitur, Fachhochschule, Hochschule	16,9	2,5	1,5	2,5	0,9	3,1	-0,2
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Bis Volks-/Hauptschul- abschluß, Mittlere Reife, Fachhochschulreife	86,0	-6,1	-4,3	-5,0	-6,2	-4,3	-5,0
Abitur, Fachhochschule, Hochschule	13,9	6,2	4,3	5,1	6,3	4,4	5,1

Tabelle 4: Differenz zum Mikrozensus beim Merkmal Stellung im Erwerbsleben (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Hauptberuflich erwerbstätig	52,2	-0,9	4,4	-1,7	-3,9	-1,8	2,8
Nicht hauptberuflich erwerbstätig	47,8	0,9	-4,4	1,7	3,9	1,8	-2,8
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Hauptberuflich erwerbstätig	52,9	4,9	0,2	-2,7	-6,5	-4,3	-1,5
Nicht hauptberuflich	47,1	-4,9	-0,2	2,7	6,5	4,3	1,5

erwerbstätig

Stellung im Erwerbsleben

Die Stellung im Erwerbsleben ist unter den analysierten Merkmalen dasjenige, bei dem Unterschiede in den Erhebungskonzepten am ehesten die Vergleichbarkeit beeinträchtigen (vgl. die Dokumentation im Anhang). In den gewählten Abgrenzungen sind die Anteile der Erwerbstätigen in den sechs Umfragen in West und Ost z.T. größer, z.T. kleiner als im Mikrozensus. Vor dem Hintergrund unterschiedlicher Erhebungskonzepte kann hier keine Aussage zum Nonresponse-Bias gemacht werden.

Berufliche Stellung

Was die berufliche Stellung betrifft, ist nach der Literatur mit einer Untererfassung der Arbeiter zu rechnen. Diese läßt sich in Westdeutschland auch bei allen Umfragen beobachten, allerdings sind die Abweichungen zumeist gering. Im Gegenzug sind in allen Surveys die Angestellten zu stark vertreten - am ausgeprägtesten im Sozialwissenschaftlichen-Bus II 1993 und in der Media-Analyse. Die beiden am besten ausgeschöpften Umfragen weisen damit auch insgesamt gesehen die größte Differenz zum Mikrozensus auf, während man den anderen Surveys bei diesem Merkmal eine relativ gute Übereinstimmung mit den amtlichen Daten attestieren kann.

Tabelle 5: Differenz zum Mikrozensus beim Merkmal Berufliche Stellung (in Prozentpunkten)*

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Arbeiter	30,2	-2,0	-0,5	-2,7	-0,7	-4,1	-2,5
Angestellte	46,5	1,7	0,9	2,3	2,3	5,4	6,1
Beamte	8,8	1,9	-0,5	1,2	-0,8	-0,1	-2,2
Selbständige	11,0	-1,0	1,6	-0,5	-0,5	-0,7	0,1
In Ausbildung	3,5	-0,7	-1,4	-0,3	-0,3	-0,4	-1,4
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Arbeiter	39,1	-3,7	-0,4	-2,6	-4,3	-6,8	-2,4
Angestellte	48,6	3,4	-4,0	0,9	-0,1	-0,1	-1,9
Beamte	1,8	-0,6	1,1	-0,4	1,0	0,8	0,9
Selbständige	6,7	1,7	4,0	2,7	3,8	5,8	3,8

In Ausbildung	3,8	-0,8	-0,7	-0,6	-0,4	0,3	-0,4
---------------	-----	------	------	------	------	-----	------

* Derzeit hauptberuflich erwerbstätige Personen.

Auch in Ostdeutschland sind Arbeiter unterrepräsentiert, am stärksten in den beiden Erhebungen des Sozialwissenschaften-Bus (-4,3 bzw. -6,8 Prozentpunkte). Dafür ist in allen Umfragen der Anteil der Selbständigen besonders hoch. In Relation zu ihrem Anteil an den Erwerbstätigen ist die Abweichung vom Mikrozensus sehr groß: im Sozialwissenschaften-Bus II 1993 z.B. ist der Anteil der Selbständigen beinahe doppelt so hoch wie im Mikrozensus.²⁰⁾ Insgesamt läßt sich auch in Ostdeutschland keine bessere Anpassung an den Mikrozensus mit einer Zunahme der Ausschöpfung feststellen.

Familienstand

In Einklang mit bisherigen Befunden läßt sich beim Merkmal Familienstand eine Unterschätzung nicht-verheirateter Personen beobachten. In West- und Ostdeutschland liegt der Anteil der Verwitweten/Geschiedenen sowie der Ledigen in allen Umfragen jeweils unter dem Wert des Mikrozensus, der Anteil der Verheirateten darüber.²¹⁾

Tabelle 6: Differenz zum Mikrozensus beim Merkmal Familienstand (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
Verheiratet	60,5	7,3	2,8	10,1	6,3	5,0	7,4
Verwitwet/Geschieden	15,2	-3,7	-1,1	-4,4	-3,7	-4,6	-2,6
Ledig	24,3	-3,6	-1,8	-5,7	-2,6	-0,4	-4,8
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
Verheiratet	64,3	10,5	3,7	12,2	8,9	2,9	8,9
Verwitwet/Geschieden	16,9	-6,2	-0,5	-4,8	-3,7	-3,0	-5,1
Ledig	18,8	-4,2	-3,3	-7,4	-5,1	0,1	-3,9

Eine geringe Differenz zum Mikrozensus besteht in beiden Teilen Deutschlands beim ALLBUS 1994, eine sehr große Abweichung beim Wohlfahrtssurvey. In dieser Umfrage liegt z.B. der Anteil der Verheirateten jeweils um mehr als zehn Prozentpunkte über dem Ergebnis des Mikrozensus. Weder in West- noch in Ostdeutschland werden die Unterschiede zum Mikrozensus mit einer Zunahme der Ausschöpfung geringer. So weist auch die Media-Analyse trotz 80 Prozent Ausschöpfung beim Familienstand starke Verzerrungen auf.

Haushaltsgröße

Auch hinsichtlich des Merkmals Haushaltsgröße ergibt sich in der Grundtendenz ein einheitliches Bild. In West- und Ostdeutschland sind wie erwartet alleinlebende Personen in allen Umfragen unterrepräsentiert. Nur geringe Unterschiede zum Mikrozensus weist wiederum der ALLBUS 1994 auf,²²⁾ die größten Abweichungen bestehen beim ALLBUS 1992 und Wohlfahrtssurvey 1993.²³⁾ Auch die Verteilungen bei diesem Merkmal zeigen, daß aus der Höhe der Ausschöpfung keine direkten Rückschlüsse auf die Verzerrung der Stichprobe getroffen werden können.

Tabelle 7: Differenz zum Mikrozensus beim Merkmal Haushaltsgröße (in Prozentpunkten)

	MZ 1993 (%)	ALLBUS 1992	ALLBUS 1994	Wohl.- Survey 1993	SoWi- Bus III 1993	SoWi- Bus II 1993	Media- Analyse 1994
West							
Ausschöpfungsquote (%)		51,9	53,2	62,9	68,0	73,0	80,2
1 Person	19,9	-7,3	-1,8	-8,4	-4,3	-5,0	-4,6
2 Personen	34,7	-1,2	0,3	-1,4	2,7	1,7	2,6
3 Personen	21,3	3,6	-0,6	2,8	0,1	1,2	0,6
4 Personen	16,8	2,7	0,9	5,0	0,7	2,3	1,3
5 und mehr Personen	7,3	2,1	1,1	1,9	0,8	-0,2	0,1
Ost							
Ausschöpfungsquote (%)		54,7	55,2	64,4	71,1	72,9	80,2
1 Person	15,9	-7,8	-2,2	-7,1	-5,9	-3,1	-5,3
2 Personen	34,6	0,7	-0,3	1,2	4,7	3,6	2,8
3 Personen	24,3	1,4	0,9	4,2	0,8	-0,8	1,7
4 Personen	20,0	3,3	1,1	1,3	-0,5	-0,1	-0,2
5 und mehr Personen	5,1	2,5	0,7	0,5	1,1	0,5	1,0

5. Fazit und Schlußbemerkungen

Als Ergebnis der durchgeführten Analysen läßt sich festhalten:

- Die analysierten Umfragen weisen bei den betrachteten soziodemographischen Merkmalen in der Regel Abweichungen in gleicher Richtung vom Mikrozensus auf.
- Die Struktur der Abweichungen stimmt im großen und ganzen mit dem Wissen über die Systematik von Ausfällen überein, so daß es naheliegt, sie als Ausdruck eines Nonresponse-Bias zu interpretieren.

- Es finden sich keine empirischen Indizien dafür, daß der Bias mit der Höhe der Ausschöpfungsquote abnimmt. Zumindest hinsichtlich der betrachteten demographischen Merkmale gibt es keine Hinweise dafür, daß besser ausgeschöpfte Umfragen geringere Stichprobenverzerrungen aufweisen als Umfragen mit niedrigerer Ausschöpfung. Die Höhe der berichteten Ausschöpfung ist somit kein guter Indikator für die Größe des Nonresponse-Bias.²⁴⁾

Wie ist dieser letzte Befund erklärbar? Hierfür kommen im Prinzip zwei Möglichkeiten in Betracht. Eine mögliche Erklärung könnte darin liegen, daß in den besser ausgeschöpften Umfragen zwar die Ausschöpfung höher, gleichzeitig aber auch die Unterschiede zwischen Teilnehmern und Nichtteilnehmern größer sind als in den schlechter ausgeschöpften Umfragen (vgl. Groves 1989: 145ff.). Ein solcher Fall tritt dann ein, wenn im Zuge der Bemühungen zur Verbesserung der Ausschöpfung lediglich solche Personen doch noch für ein Interview gewonnen werden können, die eher „untypisch“ für die gesamte Gruppe der ursprünglichen Ausfälle sind. Feldarbeitsstrategien, die lediglich darauf abzielen, die Ausschöpfungsquote insgesamt zu maximieren, bieten keine Gewähr für eine Reduzierung des Nonresponse-Bias.²⁵⁾

Plausibler scheint mir jedoch ein anderes Erklärungsmodell zu sein. Zu fragen ist nämlich, inwieweit die berichteten Ausschöpfungsergebnisse - insbesondere die hohen Werte von 70 Prozent und mehr - die Realität korrekt wiedergeben. Womöglich resultiert der fehlende Zusammenhang zwischen der Höhe der Ausschöpfung und dem Ausmaß des Nonresponse-Bias daraus, daß die betrachteten Umfragen sich de facto in der Ausschöpfung doch nicht so sehr unterscheiden, wie die berichteten Quoten es nahelegen.

Bei den analysierten Umfragen kann lediglich im ALLBUS 1994 die Ausschöpfung (weitgehend) zuverlässig berechnet werden, weil aufgrund der Ziehung einer Stichprobe von Personenadressen aus Einwohnermelderegistern vor Feldbeginn ein genau definiertes Stichprobenbrutto vorlag (zu den durchgeführten Kontrollen vgl. Koch 1995). Sämtliche anderen Umfragen dagegen wurden nach dem ADM-Design durchgeführt. Die faktische Bruttostichprobe wird hier durch die Interviewer in der zweiten und dritten Auswahlstufe gebildet. Die Zuverlässigkeit der Ausschöpfungsberechnung hängt davon ab, wie exakt die Interviewer die Vorgaben beachten. Bei diesem Verfahren ist nicht auszuschließen, daß Interviewer die strikte Einhaltung der Random-Anweisungen zugunsten der Realisierung einer möglichst großen Zahl von Interviews opfern und *undokumentiert* schwer erreichbare oder kooperationsunwillige Zielpersonen gegen leichter erreichbare und kooperative Personen *substituieren*.²⁶⁾

Ein einfaches Rechenexempel verdeutlicht, daß bereits relativ „kleine“ Manipulationen die Höhe der berichteten Ausschöpfung merklich beeinflussen können: Angenommen, ein Interviewer hätte in einem Wahlbezirk acht Adressen als Brutto vorgegeben und realisierte daraus unter Beachtung aller Regeln fünf Interviews, dann läge seine Ausschöpfung bei 62,5 Prozent. Erzielte er regelwidrig noch ein weiteres Interview mit einer Person, die nicht zu seinem Brutto gehört, könnte er seine Ausschöpfung bereits auf 75 Prozent „verbessern“!

Inwieweit ein solches Fehlverhalten bei den vorliegenden Umfragen - und in der Umfragepraxis generell - eine Rolle spielt, kann hier nicht beurteilt werden. Eine öffentliche Diskussion darüber, welche Ausschöpfungen heutzutage faktisch möglich sind, findet in der Bundesrepublik nicht oder nur in Ansätzen statt (vgl. z.B. Porst 1996). Nur vereinzelt finden sich kritische Äußerungen wie die von Sommer (1987), der generell bezweifelt, daß in allgemeinen Bevölkerungsumfragen Ausschöpfungsquoten von 70 Prozent und mehr im normalerweise gegebenen Zeit- und Kostenrahmen überhaupt noch realisierbar sind.²⁷⁾

Eine wichtige Voraussetzung für realistische Ausschöpfungsberechnungen ist, daß Umfrageinstitute und Auftraggeber Abschied nehmen von der Vorstellung, daß die Ausschöpfungsquote alleine einen Indikator für die Stichprobenqualität darstellt und die Erzielung eines bestimmten Ausschöpfungslevels quasi automatisch „Qualität“ garantiert. Es wäre wünschenswert, wenn die vorliegenden Ergebnisse dazu beitragen würden, diese Annahme zu relativieren und den „Mythos der Ausschöpfung“ (Sommer 1987) durch eine breiter angelegte Diskussion über Qualitätskriterien für Stichproben zu ersetzen.

Anmerkungen

- 1) So heißt es etwa im Rahmenschema für Werbeträgeranalysen 1988 des Zentralverbands der deutschen Werbewirtschaft auf S. 15: „Je höher die Ausschöpfung, um so besser ist die Qualität der Stichprobe“ (zit. nach Landgrebe 1992: 20).
- 2) Einen Überblick über die Verfahren zur Gewinnung von Informationen über Ausfälle geben Groves (1989: 186ff.) und Schnell (1997: 134ff.).
- 3) Dies gilt natürlich nur dann, wenn nicht andere Fehlerquellen wie etwa Meßfehler die Gültigkeit der Mikrozensus-Daten beeinträchtigen. Hartmann (1990) fand eine gute Übereinstimmung der Randverteilungen demographischer Merkmale aus dem Mikrozensus mit anderen statistischen Quellen (wie z.B. der Volkszählung oder der Hochschulstatistik).

4) Nähere Informationen zu den einzelnen Umfragen bieten: ALLBUS 1992 (Braun u.a. 1993), ALLBUS 1994 (Koch u.a. 1994), Wohlfahrtssurvey 1993 (Infratest Sozialforschung 1993), Sozialwissenschaften-Bus (GFM-GETAS o.J.), Media-Analyse (Arbeitsgemeinschaft Media-Analyse o.J.).

5) Informationen zum ADM-Stichprobenverfahren finden sich in mehreren Beiträgen in Gabler/Hoffmeyer-Zlotnik (1997). Auch die im ALLBUS 1994 erfolgte Stichprobenziehung aus Einwohnermelderegistern ist dort ausführlich beschrieben.

6) In den einzelnen Umfragen kamen dabei unterschiedliche Verfahrensweisen zum Einsatz. In der Media-Analyse wurde das sog. Adress-Random-Verfahren verwendet, bei dem eine bestimmte Zahl von Haushaltsadressen in einem getrennten Arbeitsschritt vor der Durchführung der Interviews ermittelt wird. In jedem Wahlbezirk wurden durch die Interviewer zunächst 50 Haushaltsadressen aufgenommen und im Umfrageinstitut wurden fünf Adressen für die eigentliche Befragung zufällig ausgewählt. Im Wohlfahrtssurvey und in den beiden Umfragen des Sozialwissenschaften-Bus dagegen kam das in der Praxis gebräuchlichere Random-Route-Verfahren zum Einsatz, bei dem die Haushaltsauflistung in einem Zuge mit der Realisierung der Interviews erfolgt. Die Interviewer hatten entlang des Begehungsweges in systematischer Auswahl jeweils acht bzw. neun Haushalte zu kontaktieren und in möglichst jedem dieser Haushalte ein Interview durchzuführen. Auch das - im Sprachgebrauch von Infratest - als Standard-Random bezeichnete Vorgehen beim ALLBUS 1992 entspricht dem Random-Route-Modell. Der Unterschied zum herkömmlichen Random-Route-Verfahren besteht allerdings darin, daß den Interviewern nicht mehr ein festes Brutto von zu bearbeitenden Haushalten vorgeschrieben wird, sondern lediglich eine Nettofallzahl zu realisierender Interviews, die aus einem relativ großzügig bemessenen Brutto erzielt werden kann. Im ALLBUS 1992 sollten pro Wahlbezirk sechs (West) bzw. fünf (Ost) Interviews aus maximal 15 bearbeitbaren Adressen realisiert werden.

7) In allen fünf Surveys wurde jeweils eine Person im Haushalt befragt. Eine Ausnahme bildet die Media-Analyse, bei der in großen Haushalten (mit vier und mehr Personen im Alter von 14 und mehr Jahren) zwei Befragungspersonen ausgewählt wurden.

8) Die Berechnung der Ausschöpfungsquote erfolgte in einheitlicher Weise. Bei allen sechs Umfragen wurden Ausfälle, die nicht zur Grundgesamtheit der Umfrage zählten, sowie nicht auffindbare bzw. nicht bearbeitete Adressen nicht zum einsetzbaren Brutto gezählt (sog. „stichprobenneutrale Ausfälle“). Als ausschöpfungsrelevante, nicht-stichprobenneutrale Nonresponses wurden Ausfälle aufgrund von Nichterreichbarkeit, Krankheit, Verweigerung oder mangelnder Deutschkenntnisse gewertet.

9) Es ist nicht Ziel des vorliegenden Beitrages, die Ursachen für diese unterschiedlichen Ausschöpfungsergebnisse zu analysieren. Einfache Erklärungen dafür sind kaum möglich. So ist etwa auf Basis der verfügbaren Informationen kein Zusammenhang zwischen der Ausschöpfung und der Länge der Feldzeit, der

durchschnittlichen Interviewdauer oder den Studienkosten festzustellen. Ich werde im letzten Abschnitt diese Thematik nochmals aufgreifen.

10) Man beachte, daß bei der Media-Analyse 1994 der überwiegende Teil der Interviews im Jahr 1993 durchgeführt wurde. Um die Erhebungsjahre auch für die beiden ALLBUS-Umfragen konstant zu halten, hätte man natürlich auch die Mikrozensus 1992 und 1994 heranziehen können. Ich habe dies nicht getan, weil in beiden Erhebungen Bildungsmerkmale nicht erfaßt wurden.

11) Die Ergebnisse beruhen auf den - mit dem Design-Gewicht - personenrepräsentativ gewichteten Verteilungen.

12) Unter der Annahme einfacher Zufallsstichproben wären in Westdeutschland bei den fünf sozialwissenschaftlichen Umfragen folgende Differenzen zum Mikrozensus auf dem 5-Prozent-Niveau signifikant: Bei Anteilswerten von ungefähr 10 bzw. 90 Prozent Differenzen größer als ca. 1,5 Prozentpunkte, bei Anteilswerten nahe 50 Prozent Differenzen größer als ca. 2,5 Prozentpunkte. Bei der Media-Analyse liegen die Grenzen aufgrund der höheren Fallzahl niedriger: Hier wären bereits Abweichungen von einem halben bzw. einem Prozentpunkt signifikant. In Ostdeutschland lägen die Werte aufgrund der niedrigeren Fallzahlen mit 2,0 und 3,5 Prozentpunkten bei den sozialwissenschaftlichen Umfragen und 1,0 bzw. 1,5 Prozentpunkten bei der Media-Analyse etwas höher. Diese Zahlen sind als untere Grenze zu betrachten, da die Konfidenzintervalle von Schätzern auf der Basis komplexer Stichproben, wie sie hier vorliegen, zumeist größer sind als bei einfachen Zufallsstichproben (vgl. Schnell 1997: 65ff., 272ff.).

13) Wie oben erwähnt, werden beim ADM-Design lediglich die Haushalte gleichwahrscheinlich ausgewählt. Für Auswertungen auf Personenebene ist damit designbedingt eine Gewichtung notwendig. In der Praxis ist die Anwendung dieses sog. „Transformations-Gewichts“ nicht unumstritten, weil - wie wir auch hier im folgenden noch sehen werden - das zur Gewichtung herangezogene Merkmal „Haushaltsgröße“ selbst nicht frei von Verzerrungen ist (vgl. Hartmann 1990, Hartmann/Schimpl-Neimanns 1992). Wenn Aussagen über die Qualität von realisierten Stichproben getroffen werden sollen, liegt es jedoch nahe, daß entsprechend den Vorgaben des Designs verfahren und mithin designgewichtet ausgewertet wird (vgl. allerdings Hartmann (1990: 10), der anders verfährt).

14) Eine Analyse des Teilnahmeverhaltens beim ALLBUS 1994 mit Individualdaten aus den Einwohnermelderegistern für Teilnehmer und Nichtteilnehmer zeigt, daß die Unterrepräsentierung der Frauen tatsächlich ein Nonresponse-Effekt ist (vgl. Koch 1997: 112). Die Teilnahmequote der Frauen lag in Westdeutschland mit 49,7 Prozent knapp 10 Prozentpunkte unter der der Männer mit 59,1 Prozent. Auch in Ostdeutschland war die Ausschöpfung bei Männern mit 59,2 Prozent höher als bei Frauen mit 53,7 Prozent. Diese geringere Differenz zwischen den Geschlechtern im Osten steht im Einklang mit der hier beobachtbaren geringeren Unterrepräsentierung von Frauen in Ostdeutschland.

15) Auch wenn die einzelnen Abweichungen zumeist nicht signifikant sind, bedeutet dies nicht, daß sie tatsächlich zufällig erfolgen. Nicht signifikante Abweichungen *können* zufällig sein, *müssen* es aber nicht. Es spricht stichprobentheoretisch nichts gegen ihre Interpretation, wenn es hierfür plausible Gründe - im vorliegenden Fall die Gleichförmigkeit der Abweichung, die mit den bekannten Regelmäßigkeiten des Teilnahmeverhaltens übereinstimmt - gibt. Wären die Abweichungen zufällig, dann träte die Konstellation 'sechs mal negativ' mit der Wahrscheinlichkeit von $0,5^6 = 0,015625$ auf.

16) Zu erwähnen ist, daß auch beim Merkmal Alter die vorliegenden Ergebnisse für den ALLBUS 1994 in ihrer Grundstruktur mit den Befunden auf Basis der Registerangaben (vgl. Koch 1997: 112) übereinstimmen.

17) Dies gilt erstaunlicherweise auch für die Media-Analyse, für die man nicht nur wegen der hohen Ausschöpfung, sondern auch aufgrund der höheren Fallzahlen und des dadurch niedrigeren Stichprobenfehlers eigentlich eine bessere Anpassung an den Mikrozensus als bei den anderen Umfragen erwarten würde.

18) Darauf hinzuweisen ist, daß die in der Tabelle berichteten Abweichungen vom Mikrozensus die tatsächliche Untererfassung niedriger Bildungsabschlüsse vermutlich sogar noch unterschätzen. Da im Mikrozensus 1993 die Fragen zum Bildungsabschluß nicht der Auskunftspflicht unterlagen, weisen diese einen Item-Nonresponse von ca. 10 bis 15 Prozent auf. Analysen zeigen, daß dieser Item-Nonresponse selektiv erfolgt: insbesondere Bildungsabschlüsse unterer Qualifikationsniveaus werden im Mikrozensus unterschätzt (vgl. Riede/Emmerling 1994).

19) Hauptgrund hierfür ist die unterschiedliche Erfassung des Abschlusses „Polytechnische Oberschule“ der ehemaligen DDR.

20) Es kann nicht völlig ausgeschlossen werden, daß für dieses Ergebnis neben einem Nonresponse-Effekt auch Unterschiede im Antwortverhalten zwischen Mikrozensus einerseits und den Umfragen andererseits eine Rolle spielen. Im Mikrozensus wurde bei der betreffenden Frage die Antwortkategorie der Selbständigen nicht weiter differenziert, in allen sechs Umfragen dagegen wurde explizit zwischen selbständigen Landwirten, (akademischen) freien Berufen und sonstigen Selbständigen unterschieden (vgl. die Rekodierungen im Anhang). Da die Abfrage in West- und Ostdeutschland in gleicher Weise erfolgte, müßte eine Erklärung auf der Basis unterschiedlichen Antwortverhaltens allerdings annehmen, daß eine differenzierte Abfrage lediglich in Ostdeutschland zu systematisch anderen Einstufungen der Befragten führt als eine summarische.

21) Lediglich beim Sozialwissenschaften-Bus II 1993 stimmt der Anteil der Ledigen in Ostdeutschland nahezu exakt mit dem Ergebnis des Mikrozensus überein.

22) Erwähnenswert ist in diesem Zusammenhang, daß auch bei einer internen Validierung der Stichprobenqualität durch Sodeur (1997: 80ff.) der auf Basis einer

Einwohnermelderegister-Stichprobe durchgeführte ALLBUS 1994 sehr gut abschneidet. In den nach dem ADM-Verfahren realisierten ALLBUS-Erhebungen der Jahre 1988 bis 1992 (mit zum Teil weit höherer berichteter Ausschöpfung, Anm. Achim Koch) stellt Sodeur dagegen Abweichungen von einer reinen Wahrscheinlichkeitsauswahl fest, die nach seiner Vermutung mit dem größeren Spielraum der Interviewer bei der Auswahl der Befragten bei diesem Verfahren zusammenhängen.

23) Das Muster der im Gegenzug überrepräsentierten Haushaltsgrößen ist unterschiedlich. Während beim ALLBUS 1992 und Wohlfahrtssurvey 1993 in West- und Ostdeutschland Personen aus Haushalten mit drei und mehr Personen zu stark vertreten sind, ist bei den Umfragen des Sozialwissenschaften-Bus und der Media-Analyse im Westen die Überrepräsentierung von Personen aus Zwei- und Vierpersonenhaushalten besonders auffällig, im Osten die Überrepräsentierung von Personen aus Zwei- sowie Fünf(undmehr)personenhaushalten.

24) Ob dies auch für andere, hier nicht untersuchte Merkmale zutrifft, kann nicht beurteilt werden. Wie in der Einleitung ausgeführt wurde, ist der Nonresponse-Bias eine merkmalspezifische Größe. Verallgemeinerungen der Ergebnisse über die betrachteten Merkmale hinaus sind damit weder möglich noch beabsichtigt.

25) Für die Gestaltung der Feldarbeit ist es deshalb notwendig, über Mittel und Wege nachzudenken, wie man nicht nur die Ausschöpfung erhöhen, sondern gleichzeitig auch die Struktur der Stichprobe verbessern kann. Dies ist nur durch den *gezielten* Einsatz ausschöpfungssteigernder Maßnahmen (wie zusätzlicher callbacks, incentives, etc.) möglich, die dazu beitragen, bestehende Ungleichgewichte in den Teilnahmewahrscheinlichkeiten verschiedener Bevölkerungsgruppen abzubauen.

26) Ähnlich äußert sich Schnell (1997: 58f.). Auch in einer „offiziellen“ Darstellung des ADM-Designs wird darauf hingewiesen, daß zumindest beim ADM-Design ohne Adressenvorlauf die Gefahr besteht, daß Interviewer „unbequeme“ Adressen übergehen (vgl. von der Heyde/Löffler 1993: 52). In den Routinekontrollen der Umfrageinstitute können solche Abweichungen vom Stichprobenplan durch die Befragung einer falschen Person - im Unterschied etwa zur Totalfälschung eines Interviews - nicht entdeckt werden.

Erwähnenswert ist in diesem Zusammenhang im übrigen, daß insbesondere das in Deutschland bei face-to-face-Umfragen übliche Entlohnungssystem keinen Anreiz für die Interviewer bietet, sich mit „schwierigen“ Fällen zu befassen: honoriert werden lediglich realisierte Interviews, und zwar zu einem festen „Stückpreis“.

27) Nach Sommer (1987) wäre es im übrigen falsch, geschönte Ausschöpfungsergebnisse einseitig den Interviewern anzulasten. Nach seiner Ansicht liegt der Fehler im System, da Abweichungen von den Random-Anweisungen von den Umfrageinstituten stillschweigend erwartet und toleriert würden.

Literatur

- Arbeitsgemeinschaft Media-Analyse, o. J.: Dokumentation MA 94. Frankfurt/M.
- Braun, M./Eilinghoff, C./Gabler, S./Wiedenbeck, M., 1993: Methodenbericht zur „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1992. ZUMA-Arbeitsbericht 93/01.
- Burton, J., 1996: Public Attitudes and Responses to Survey Research: A Review of the Literature. Unpubl. paper. University of Essex.
- Esser, H./Grohmann, H./Müller, W./Schäffer, K.-A., 1989: Mikrozensus im Wandel. Untersuchungen und Empfehlungen zur inhaltlichen und methodischen Gestaltung. Bericht des wissenschaftlichen Beirates für Mikrozensus und Volkszählung. Frankfurt/M.
- Gabler, S./Hoffmeyer-Zlotnik, J. H. P. (Hrsg.), 1997: Stichproben in der Umfragepraxis. Opladen: Westdeutscher Verlag.
- GFM-GETAS, o. J.: Sozialwissenschaften-Bus II/III 1993 - Methodenbericht zur technischen Organisation und Durchführung. Hamburg.
- Goyder, J., 1987: The Silent Minority. Nonrespondents on Sample Surveys. Cambridge: Polity Press.
- Groves, R. M., 1989: Survey Errors and Survey Costs. New York: John Wiley & Sons.
- Groves, R. M./Couper, M. P., 1992: Correlates of Nonresponse in Personal Visit Surveys. S. 102-111 in: Proceedings of the American Statistical Association, Survey Research Methods Section.
- Hartmann, P. H., 1990: Wie repräsentativ sind Bevölkerungsumfragen? Ein Vergleich des ALLBUS und des Mikrozensus. ZUMA-Nachrichten 26: 7-30.
- Hartmann, P. H./Schimpl-Neimanns, B., 1992: Sind Sozialstrukturanalysen mit Umfragedaten möglich? Analysen zur Repräsentativität einer Sozialforschungsumfrage. Kölner Zeitschrift für Soziologie und Sozialpsychologie 44: 315-340.
- Infratest Sozialforschung, 1993: Wohlfahrtssurvey 1993 - Methodenbericht. München.
- Koch, A., 1995: Gefälschte Interviews: Ergebnisse der Interviewerkontrolle beim ALLBUS 1994. ZUMA-Nachrichten 36: 89-105.
- Koch, A., 1997: Teilnahmeverhalten beim ALLBUS 1994. Soziodemographische Determinanten von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft. Kölner Zeitschrift für Soziologie und Sozialpsychologie 49: 98-122.
- Koch, A./Gabler, S./Braun, M., 1994: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 1994. ZUMA-Arbeitsbericht 94/11.
- Landgrebe, K. P., 1992: Ausschöpfungen. Planung und Analyse 19(2): 19-22.

Lüttinger, P./Riede, Th., 1997: Der Mikrozensus: Amtliche Daten für die Sozialforschung. ZUMA-Nachrichten 41: 19-43.

Porst, R., 1996: Ausschöpfungen bei sozialwissenschaftlichen Umfragen. Die Sicht der Institute. ZUMA-Arbeitsbericht 96/07.

Riede, Th./Emmerling, D., 1994: Analysen zur Freiwilligkeit der Auskunftserteilung im Mikrozensus. Sind Stichprobenergebnisse bei freiwilliger Auskunftserteilung verzerrt? Wirtschaft und Statistik 9: 733-742.

Schnell, R., 1997: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske + Budrich.

Sodeur, W., 1997: Interne Kriterien zur Beurteilung von Wahrscheinlichkeitsauswahlen. ZA-Information 41: 58-82.

Sommer, R., 1987: Der Mythos der Ausschöpfung. Planung und Analyse 14: 300-301.

von der Heyde, Ch./Löffler, U., 1993: Die ADM-Stichprobe. Planung und Analyse 20(5): 49-53.

Anhang: Definition der Kategorien

1. Schulabschluß

In den Auswertungen für Ostdeutschland wurden Kategorie 1 und 2 zusammengefaßt.

Mikrozensus

1. Bis Volks-/Hauptschulabschluß:	-kein Schulabschluß -Haupt- oder Volksschulabschluß
2. Mittlere Reife, Fachhochschulreife:	-Realschulabschluß (Mittlere Reife) -polytechnische Oberschule der DDR -Fachhochschulreife
3. Abitur, Fachhochschul-/Hochschulabschluß	-Hochschulreife (Abitur) -Fachhochschule, Ingenieurschule -Hochschule

ALLBUS, Wohlfahrtssurvey

1. Bis Volks-/Hauptschulabschluß:	-Schule beendet ohne Abschluß -Volks-/Hauptschulabschluß, Polytechnische Oberschule mit Abschluß 8. oder 9. Klasse
2. Mittlere Reife, Fachhochschulreife:	-Mittlere Reife, Realschulabschluß, Polytechnische Oberschule mit Abschluß 10. Klasse -Fachhochschulreife (Abschl. einer Fachobersch. etc.)
3. Abitur, Fachhochschul-/Hochschulabschluß	-Abitur, Erweiterte Oberschule mit Abschluß 12. Klasse (Hochschulreife) -Fachhochschulabschluß (auch Abschluß einer Ingenieurschule) -Hochschulabschluß

Schüler und Personen mit „anderem Schulabschluß“ sind in der Prozentuierung nicht berücksichtigt.

Sozialwissenschaften-Bus

1. Bis Volks-/Hauptschulabschluß:	-von der Schule abgegangen ohne Abschluß, vor Erreichen einer 9., 10., oder 11. Klasse -Volks-, Hauptschulabschluß
2. Mittlere Reife, Fachhochschulreife:	-Mittlere Reife, Realschulabschluß, Fachschulreife -Polytechnische Oberschule mit 10. Klasse Abschluß (vor 1965: 8. Klasse Abschluß) -Fachhochschulreife, Fachgebundene Hochschulreife, Abschluß einer Fachoberschule
3. Abitur, Fachhochschul-/Hochschulabschluß	-Abitur, allgemeine Hochschulreife, Erweiterte Oberschule (EOS)

- | |
|--|
| -Fachhochschulabschluß (auch Ingenieurschulabschluß) |
| -Hochschulabschluß ohne Diplom |
| -Hochschulabschluß mit Diplom/Promotion |

Schüler und Personen mit „anderem Schulabschluß“ sind in der Prozentuierung nicht berücksichtigt.

Media-Analyse

1. Bis Volks-/Hauptschulabschluß:	-Volksschule/Grundschule, Hauptschule
2. Mittlere Reife, Fachhochschulreife:	-Mittlere Reife bzw. weiterführende Schule ohne Abitur
3. Abitur, Fachhochschul-/Hochschulabschluß:	-Abitur, Hochschulreife -Studium (Universität, Hochschule, Fachhochschule, Akademie, Polytechnikum)

2. Stellung im Erwerbsleben

Mikrozensus

1. Hauptberuflich erwerbstätig:	-Personen, deren normalerweise geleistete Arbeitszeit je Woche 17 oder mehr Stunden beträgt
2. Nicht hauptberuflich erwerbstätig:	-Nicht erwerbstätige Personen -Personen, deren normalerweise geleistete Arbeitszeit je Woche 16 oder weniger Stunden beträgt

ALLBUS

1. Hauptberuflich erwerbstätig:	-Hauptberufliche Erwerbstätigkeit, ganztags -Hauptberufliche Erwerbstätigkeit, halbtags -In Kurzarbeit
2. Nicht hauptberuflich erwerbstätig:	-Nicht erwerbstätig -Nebenher erwerbstätig

Wohlfahrtssurvey

1. Hauptberuflich erwerbstätig:	-Hauptberuflich erwerbst., ganztags (35 Std. und mehr) -Hauptberuflich erwerbst., halbtags (unter 35 Std.)
2. Nicht hauptberuflich erwerbstätig:	-Nebenher erwerbstätig, regelmäßig (unter 20 Std.) -Nebenher erwerbstätig, unregelmäßig -Nicht erwerbstätig (auch arbeitslos) -Wehr- oder Zivildienst

Sozialwissenschaften-Bus

1. Hauptberuflich erwerbstätig:	-Vollzeit erwerbstätig, Wochen-Arbeitszeit West: 35 Std. und mehr, Ost: 40 Std. und mehr -Teilzeit erwerbstätig, Wochen-Arbeitszeit West: 17 bis 34 Std., Ost: 20 bis 39 Std.
---------------------------------	--

	-Auszubildende/Lehrlinge
2. Nicht hauptberuflich erwerbstätig:	-Nicht erwerbstätig -Teilzeit oder stundenweise erwerbstätig, Wochen- Arbeitszeit West: unter 17 Std., Ost: unter 20 Std. -(Vollzeit-) Schüler

Media-Analyse

1. Hauptberuflich erwerbstätig:	-Voll berufstätig -Teilweise berufstätig -In Ausbildung: Lehrling
2. Nicht hauptberuflich erwerbstätig:	-Vorübergehend arbeitslos -Rentner, Pensionär (nicht mehr voll berufstätig) -In Ausbildung: Schüler, Student -Nicht berufstätig

3. Berufliche Stellung**Mikrozensus**

1. Arbeiter:	-Arbeiter, Heimarbeiter
2. Angestellte:	-Angestellte, Zivildienstleistende
3. Beamte:	-Beamte, Richter, Polizei, BGS, Soldaten
4. Selbständige:	-Selbständige, mit und ohne Beschäftigte -Mithelfende Familienangehörige
5. In Ausbildung:	-Kaufmännisch-technische Auszubildende -Gewerbliche Auszubildende

ALLBUS

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellter
3. Beamte:	-Beamter/Richter/Berufssoldat
4. Selbständige:	-Selbständiger Landwirt (auch Genossenschaftsbauer) -Akademischer freier Beruf -Selbständiger in Handel, Gewerbe, Industrie, Dienstleistung u.a. -Mithelfender Familienangehöriger
5. In Ausbildung:	-In Ausbildung

Wohlfahrtssurvey

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellte
3. Beamte:	-Beamte/Richter/Berufssoldaten
4. Selbständige:	-Selbständige Landwirte -Akademische freie Berufe -Sonstige Selbständige (außerhalb der Landwirtschaft) -Mithelfende Familienangehörige
5. In Ausbildung:	-In Ausbildung

Sozialwissenschaften-Bus

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellter
3. Beamte:	-Beamter/Richter/Berufssoldat
4. Selbständige:	-Selbständiger Landwirt/Selbständiger Bauer oder Gärtner (auch Genossenschaftsbauer) -Akademischer freier Beruf -Selbständiger in Handel, Gewerbe, Industrie, Dienstleistung u.a. (auch PGH) -Mithelfende Familienangehörige
5. In Ausbildung:	-In Ausbildung

Media-Analyse

1. Arbeiter:	-Arbeiter
2. Angestellte:	-Angestellte
3. Beamte:	-Beamte
4. Selbständige:	-Selbständige -Freie Berufe -Selbständige Landwirte (Vollerwerbsbetrieb)
5. In Ausbildung:	-In Ausbildung: Lehrling